凸計画問題における内点法と双対幾何(非線形可積分系による応用解析)

小原, 敦美

数理解析研究所講究録
第 889 巻 1994 年 19-25

凸計画問題における内点法と双対幾何

大阪大学基礎工学部　小原 敦美　(Atsumi Ohara)

1　はじめに

Karmarkar が線形計画問題において多項式オーダーの手間の性能を持つ内点法アルゴリズムを発表して以来，アルゴリズムの改良や理論的解析が多くの研究者によってなされている．特に Riemann 幾何，情報幾何といった観点からの研究として [Bayer & Lagarias 89, Karmarkar 90, 田辺，土谷 88]，可積分系という観点からの研究として [Brockett 88] [Nakamura 94] が上げられる．

一方，近年の研究により内点法は線形計画問題から凸計画問題へと適用範囲を広げている [Nesterov & Nemirovsky 94, 吉瀬 94, Jarre 92]．特に線形行列不等式問題と呼ばれるある凸計画問題は，Lyapunov 関数，エネルギーといった正値性との関連からシステム制御理論と基礎的な部分で深いつながりがあり [Boyd et al. 94, 小原，杉江 94]，実際の応用分野における制御系設計問題で必要となるような大規模な問題に対して内点法のような効率の良いアルゴリズムの研究が不可欠とされている．

このメモでは，凸計画法の内点法でも Legendre 変換，∇^*-測地線といった双対幾何の構造が自然に現れ，これらを用いて内点法の中でもパス追跡法と呼ばれる一連のアルゴリズムが統一的に理解できることを示す．これは [田辺，土谷 88] の線形計画法における結果の簡単な拡張とみなすことができる．

2　準備：双対幾何構造の導入

$\mathcal{M}$ を $\mathbf{R}^n$ の開領域，(θ^i) を $\mathbf{R}^n$ の座標系とする．また，$\psi(\theta)$ を $\mathcal{M}$ 上のなめらかな狭義凸関数とする．$\psi(\theta)$ をポテンシャル関数として，$\mathcal{M}$ 上の双対幾何構造が次のように導かれる [Amari 85, Fujiwara 93]．

$\psi(\theta)$ は狭義凸であるからその Hesse 行列は正定でこれを $\mathcal{M}$ 上の Riemann 計量 g とする：

$$g_{ij}(\theta) = \partial_i\partial_j\psi(\theta), \text{ただし}, \partial_i := \frac{\partial}{\partial\theta^i} \tag{2.1}$$

次に

$$T_{ijk}(\theta) := \partial_i\partial_j\partial_k\psi(\theta) \tag{2.2}$$

として，実数αによって定まる $\mathcal{M}$ 上の接続 $\nabla^{(\alpha)}$の係数を

$$\Gamma^{(\alpha)}_{ijk}(\theta) := g(\nabla_{\partial_i}\partial_j, \partial_k) = [ij;k] - \frac{\alpha}{2}T_{ijk}(\theta) \tag{2.3}$$

と定義する．ただし，$[ij;k]$ は Riemann 接続の係数

$$[ij;k] := \frac{1}{2}(\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij})$$

である．特に$\alpha = \pm 1$ のときの接続をそれぞれ $\nabla = \nabla^{(1)}, \nabla^* = \nabla^{(-1)}$としよう．対応する接続の係数の$\theta$座標系での表現は

$$\Gamma_{ijk}(\theta) = [ij;k] - \frac{1}{2}T_{ijk}(\theta) = 0, \quad \Gamma^*_{ijk}(\theta) = [ij;k] + \frac{1}{2}T_{ijk}(\theta) = T_{ijk}(\theta) = \partial_i g_{jk}(\theta) \quad (2.4)$$

となる．$\Gamma_{ijk}(\theta) = 0$ から，座標系θは ∇-アフィン座標系とよばれる．また，上式より任意の $\mathcal{M}$ 上のベクトル場 A, B, Cに対して

$$\partial_i g_{jk} = \Gamma_{ijk} + \Gamma^*_{ikj} \Leftrightarrow Ag(B,C) = g(\nabla_A B, C) + g(B, \nabla^*_A C) \quad (2.5)$$

が成り立つので，∇ と ∇^*は gに対して互いに双対な接続と呼ばれる．以上のように $\mathcal{M}$ 上の凸ポテンシャル$\psi(\theta)$ から，双対幾何の構造 $(\mathcal{M}, g, \nabla, \nabla^*)$ が導かれる．

さて，双対な接続が凸ポテンシャル$\psi(\theta)$ から (2.4) のように導かれた場合，∇ と ∇^*に対して $\mathcal{M}$ の捩率，曲率はともに 0 となる．これを $\mathcal{M}$ は双対平坦であると呼ぶ．

双対平坦な $\mathcal{M}$ 上には Legendre 変換により，

$$\eta_i := \partial_i \psi(\theta), \quad \phi(\eta) := \sup_{\theta \in \mathcal{M}} \{\theta^i \eta_i - \psi(\theta)\} \quad (2.6)$$

と定義される座標系 (η_i) とηに関する狭義凸関数$\phi(\eta)$ を導入することができる．Riemann 計量 g，双対接続 ∇,∇^*のこの座標系での成分（これを上付き添え字でかく）は$\partial^i := \frac{\partial}{\partial \eta_i}$として，

$$g^{ij}(\eta) = \partial^i \partial^j \phi(\eta) \quad (2.7)$$

$$\Gamma^{ijk}(\eta) = \partial^i \partial^j \partial^k \phi(\eta), \quad \Gamma^{*ijk}(\eta) = 0 \quad (2.8)$$

となる．ただし，$\Gamma^{*ijk}(\eta) = 0$ からηは ∇^*-アフィン座標系と呼ばれる．また，θとηの変換の Jacobi 行列は gのそのもの，すなわち

$$\partial_i = g_{ij} \partial^j, \quad \partial^i = g^{ij} \partial_j \quad (2.9)$$

であり，$(g_{ij}(\theta))$ と $(g^{ij}(\eta))$ は同一の点 $p = p(\theta) = p(\eta)$ では逆行列の関係にある．

最後に接続 ∇ に関する測地線は，θ座標，η座標でそれぞれ

$$\ddot{\theta}^i(t) = 0, \quad g^{ij}\ddot{\eta}_j(t) + \Gamma^{ijk}\dot{\eta}_j(t)\dot{\eta}_k(t) = 0$$

で表され，同様に接続 ∇^*に関する測地線は

$$g_{ij}\ddot{\theta}^j(t) + \Gamma^*_{ijk}\dot{\theta}^j(t)\dot{\theta}^k(t) = 0, \quad \ddot{\eta}_i(t) = 0$$

となることに注意する．

3 凸計画問題の内点法

次のような凸関数 $f_i(x)$ によって定められる $\mathbf{R}^n$ の凸集合 $\mathcal{M}$ 上の線形関数 $c^T x$ を最小化する凸計画問題を考える：

$$\min c^T x, \quad x \in \mathcal{M} \subset \mathbf{R}^n \tag{3.1}$$
$$\mathcal{M} = \{x | f_i(x) \le 0, i = 1, \ldots, m\}, \quad f_i(x) : convex, \quad i = 0, \ldots m$$

一般的な凸計画問題

$$\min f_0(x), \quad x \in \mathcal{M} \subset \mathbf{R}^n \tag{3.2}$$
$$\mathcal{M} = \{x | f_i(x) \le 0, i = 1, \ldots, m\}, \quad f_i(x) : convex, \quad i = 0, \ldots m$$

は新しい変数 x^{n+1} と新しい関数

$$f_{m+1}(x, x^{n+1}) := f_0(x) - x^{n+1} \tag{3.3}$$

を用いて

$$\min x^{n+1} = c^T \tilde{x}, \quad \tilde{x} \in \tilde{\mathcal{M}} \subset \mathbf{R}^{n+1}$$
$$c = [\, 0 \ \ldots \ 1 \,], \quad \tilde{x} = [\, x \ x^{n+1} \,]$$
$$\tilde{\mathcal{M}} = \{\tilde{x} | f_i(x) \le 0, i = 1, \ldots, m+1\}, \quad f_i(x) : convex, \quad i = 0, \ldots m+1$$

と (3.1) の形に帰着できるので，問題 (3.1) は凸計画問題の一つの標準的な形といえる．

ここでは $\mathcal{M}$ は有界としておく．この仮定がなくとも，多くの意味のある問題ならば十分大きな数λに対して $c^T x \le \lambda$という制約をつけることで，$\mathcal{M}$ を有界とする事ができると思われる．

ここで

$$\psi(x) := -\sum_{i=1}^{m} \log(-f_i(x)) \tag{3.4}$$

としよう．$\psi(x)$ は 1) int$\mathcal{M}$ 上で狭義凸，2)$\psi(x) \to \infty, (x \to \partial\mathcal{M})$ となっていることに注意しておく．$\psi(x)$ は $\mathcal{M}$ の log barrier 関数，また log barrier 関数を最小にする点は $\mathcal{M}$ の解析的中心と呼ばれ（これを x_{AC}と記そう），凸計画問題の内点法では重要な働きをする．

内点法には大きく分けるとパス追跡法 (Path Following Method) とポテンシャル減少法 (Potential Reduction Method) がある [Nesterov & Nemirovsky 94] が，ここではパス追跡法と呼ばれる凸計画問題に対する一連の内点法の考え方を問題 (3.1) について簡単に述べる．

i) バリア法 (Barrier Method)

問題の最適解を求めるためにバリア関数と目的関数の重みつき和

$$\Psi_\mu(x) := \mu c^T x + \psi(x) \to \min \tag{3.5}$$

を考える．ここでμは重みパラメータである．$x^*(\mu)$ を$\Psi_\mu(x)$ の最適解，x^*を問題 (3.1) の最適解とすると

$$x^*(\mu) \to x^* \quad (\mu \to +\infty)$$

となる．実際にはμを $\{\mu_i\}$ と適当に離散化し，$x^*(\mu_i)$ の近似解を Newton 法で見つけて $x^*(\mu)$ を追跡する．

ii) 中心法 (Method of Centers)
適当な定数$\zeta > 0$ を用いて

$$\Psi_t(x) := -\zeta \log(t - c^T x) + \psi(x) \to \min \tag{3.6}$$

を考える．$x^*(t)$ を$\Psi_t(x)$ の最適解，t^*を問題 (3.1) の最適解 x^*における値（最小値）とすると

$$x^*(t) \to x^* \quad (t \to t^*)$$

となる．実際には i) と同様に $t = t_i$を離散的に減らしながら，$x^*(t_i)$ の近似解を Newton 法で見つけて $x^*(t)$ を追跡する．

iii) 主平行曲線法 (Primal Parallel Trajectories Method)
超平面 $c^T x = \tau$上でバリア関数$\psi(x)$ の最小化

$$\psi(x) \to \min, \quad \text{s.t.} \quad c^T x = \tau$$

を考える．その最適解を $x^*(\tau)$ とすると，$x^*(t)$ と同様

$$x^*(\tau) \to x^* \quad (\tau \to t^*)$$

となる．実際には$\tau = \tau_i$を減らしながら，$x^*(\tau_i)$ の超平面上での近似解を Newton 法で見つけて $x^*(\tau)$ を追跡する．

以上の３つの方法におけるそれぞれの追跡すべき曲線 $x^*(\mu), x^*(t), x^*(\tau)$ は，径数が異なるものの同一の曲線であることが次のようにしてわかる．

i) では$\Psi_\mu(x)$ が狭義凸であることから，任意の$\mu > 0$ に対して $x^*(\mu)$ は

$$\mu c^T + \mathrm{grad}\psi(x^*(\mu)) = 0 \tag{3.7}$$

の解として一意に定まる．同様に$\Psi_t(x)$ の狭義凸性より，ii) では任意の $t > t^*$に対して $x^*(t)$ は

$$\frac{\zeta}{t - c^T x} + \mathrm{grad}\Psi(x^*(t)) = 0 \tag{3.8}$$

の解として一意に定まる．任意の$\mu > 0$ に対して，(3.7) から定まる $x^*(\mu)$ を用いて，t と $x^*(t)$ を

$$t := \frac{\zeta}{\mu} + c^T x^*(\mu), \quad x^*(t) := x^*(\mu) \tag{3.9}$$

とおくと，t と $x^*(t)$ は (3.8) を満たすことがわかる．従って $x^*(\mu)$ と $x^*(t)$ が同じ曲線であることがわかる．

さらに，iii) でも$\psi(x)$ の狭義凸性より，$c^T x = \tau$上で$\psi(x)$ を最小にする点 $x^*(\tau)$ は唯一である．Lagrange 関数 $L(x,\lambda) := \psi(x) + \lambda(c^T x - t)$ を用いると

$$\lambda c^T + \mathrm{grad}\psi(x^*(\lambda)) = 0, \tag{3.10}$$

の解としてλをパラメータとして一意に定まる $x^*(\lambda)$ で

$$c^T x^*(\lambda) - \tau = 0 \tag{3.11}$$

を満たすものが $x^*(\tau) = x^*(\lambda(\tau))$ である．よって，上と同様な議論で $x^*(\mu)$ と $x^*(\lambda)$ が（従って $x^*(\tau)$ も）同じ曲線であることがわかる．

結局，上の3つのパス追跡法はいずれも中心曲線 $x^*(\mu)$ を離散的に追いかける方法といって良い．しかも，Newton 法を用いるので，バリア関数$\psi(x)$ がある意味で二次関数に近ければ（これを [Jarre 92] では relatively Lipschiz 条件，[Nesterov & Nemirovsky 94] では self concordant 条件として明確にしている)，多項式オーダーの手間で解けることが保証されている．

さて，ここで $\mathcal{M}$ 上にバリア関数$\psi(x)$ から導かれる双対幾何構造を用いて中心曲線が ∇-測地線になることを示そう．

$$\Phi_\mu(x) := \mu c^T x + \psi(x), \quad c^T = [\, c_1 \quad c_2 \quad \ldots \quad c_n \,] \tag{3.12}$$

とする．$\Phi_\mu(x)$ も狭義凸となるから，$x^*(\mu)$ は極値条件

$$\frac{\partial \Phi_\mu(x)}{\partial x^i} = \mu c_i + \frac{\partial \psi(x)}{\partial x^i} = 0, i = 1, \ldots, n \tag{3.13}$$

の唯一の解であり，双対座標を用いると

$$y_i(\mu) = -c_i \mu \tag{3.14}$$

という直線，すなわち ∇^*-測地線となる．これは原点を初期点とする一階の微分方程式

$$\frac{dy_i}{d\mu} = -c_i \tag{3.15}$$

の解であり，Legendre 変換でもとの座標 x に戻すと $x(\mu)$ は

$$\frac{dx^i}{d\mu} = -g^{ij} c_j = -g^{ij} \partial_j (c^T x) \tag{3.16}$$

という Riemann 計量のもとでの勾配系の解となる．初期点は $y = \mathrm{grad}\psi(x) = 0$ となる点，すなわち $\mathcal{M}$ の解析的中心 x_{AC}である．あるいは ∇^*-測地線としての微分方程式

$$g_{ij} \frac{d^2 x^j}{d\mu^2} + \Gamma^*_{ijk} \frac{dx^j}{d\mu} \frac{dx^k}{d\mu} = 0 \tag{3.17}$$

で初期点が解析的中心，初期速度方向が$-g^{ij}(x_{\mathrm{AC}}) c_i$の解としても与えられる．

4 おわりに

凸計画問題における内点法で主要な働きをする中心曲線が，バリア関数から導かれる双対幾何の測地線として特徴づけられることを述べた．パス追跡法はインプリメントが簡単で汎用性もあり，もちろんその多項式性によって最悪の問題についての性能は保証されている．しかし，中心曲線を追跡するという性質から実際的な問題に対しての性能は必ずしも良くないということが指摘されており [Nesterov & Nemirovsky 94, 吉瀬 94]，線形行列不等式問題における簡単な数値実験でもステップサイズがあまり大きくないという事実が確認できる．

アルゴリズムの高速化は工学の現場では切実な問題である．幸いにも本共同研究では，加速法や保存量を考慮した数値積分法，双対座標（Legendre 変換）の役割が可積分系の立場から研究されている．このような可積分性，幾何学といった解析的なアプローチからの知見，道具が，実用的なアルゴリズムの改良あるいは新しいアルゴリズムの開発につながることを期待する．

参考文献

[Bayer & Lagarias 89] D. A. Bayer and J. C. Lagarias, The Nonlinear Geometry of Linear Programming I, *Trans. Amer. Math. Soc.*, 314, 499-526 (1989), II, *Trans. Amer. Math. Soc.*, 314, 527-581 (1989).

[Karmarkar 90] N. Karmarkar, Riemannian Geometry Underlying Interior-Point Methods for Linear Programming, *Contemporary Mathematics*, 114, 51-75 (1990)

[田辺，土谷 88] 田辺，土谷，線形計画の新しい幾何学，数理科学，303, 32-37 (1988)

[Brockett 88] R. W. Brockett, Dynamical Systems That Sort Lists and Solve Linear Programming Problems, *Proc. 27th IEEE CDC*, 779-803 (1988)

[Nakamura 94] Y. Nakamura, Lax Pair and Fixed Point Analysis of Karmarkar's Scaling Trajectory for Linear Programming, *Japan J. Indust. Appl. Math.*, 11, 1-9 (1994)

[Nesterov & Nemirovsky 94] Yu. Nesterov and A. Nemirovsky, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM (1994).

[吉瀬 94] 吉瀬，凸計画問題に対する最適化手法 -内点法と解析的中心-, システム/制御/情報，Vol.38, No.3, 155-160 (1994)

[Jarre 92] F. Jarre, Interior-Point Methods for Convex Programming, *Applied Mathematics and Opitimization*, Vol.26, 287-311 (1992)

[Boyd et al. 94] S. Boyd, L. El Ghaoui, E. Feron, V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory,* SIAM (1994)

[小原，杉江 94] 小原，杉江，凸最適化を用いた制御系設計，システム/制御/情報，Vol.38, No.3, 139-146 (1994)

[Amari 85] S. Amari, *Differential-Geometrical Methods in Statistics* Springer-Verlag, (1985).

[Fujiwara 93] A. Fujiwara, Dynamical Systems on Statistical Models, *RIMS Kokyuroku*, 822, 32-42 (1993)